

PATENT APPLICATION

Information Processing System, Storage System, Storage Device Control Apparatus and Program

Inventors: **Noboru FURUUMI**
Citizenship: Japan

Yoshihiro ASAKA
Citizenship: Japan

Junichi MUTO
Citizenship: Japan

Assignee: **Hitachi, Ltd.**
6, Kanda Surugadai 4-chome
Chiyoda-ku, Tokyo, Japan
Incorporation: Japan

Entity: Large

**INFORMATION PROCESSING SYSTEM, STORAGE SYSTEM,
STORAGE DEVICE CONTROL APPARATUS AND PROGRAM**

CROSS-REFERENCE TO RELATED APPLICATIONS

5 This application claims priority to Japanese Patent Application No. 2002-343436 filed on November 27, 2002, the content of which is incorporated herein by reference.

BACKGROUND OF THE INVENTION

10 **1. Field of the Invention**

The present invention relates to an information processing system, a storage system, a storage device controller and a program.

15 **2. Description of the Related Arts**

Recently, a technique called remote copying has been developed in order to prevent losing of data stored in a computer system and to enable an immediate recovery of the system even when a disaster such as a great earthquake occurs. The remote copying is a technique with which a duplicate of data is stored in a storage apparatus of a back-up computer system placed on a remote site (see, e.g., Japanese Patent Application Laid-open (Kokai) Pub. No. 2001-306414).

25 Thereby, it is possible to carry out tasks using a computer system on the remote site even if the operation of the computer system on the main site is terminated

when a disaster occurs.

To realize the remote copying described above, it is necessary to set up a data communication path between the main site and the remote site. In a conventional remote copying, when the two storage apparatuses respectively try to set up a data communication path such that each storage apparatus is the sender of data respectively, it is impossible to form the data communication path of each storage apparatus on a one same communication path physically.

Furthermore, when the computer system on the main site has been recovered, it is necessary to update the data on the main site by duplicating the data updated on the remote site after a failure has occurred on the main site. Therefore, settings need to be made such that the data is transmitted in the direction from the remote site to the main site but much time and labor are necessary to make such settings.

20

SUMMARY OF THE INVENTION

The present invention was conceived based on the above and other viewpoints. One object of the invention is to provide an information processing system, a storage system, a storage device controller and a program that utilize effectively a communication path for carrying out remote copying and enable facilitating the recovery work of the remote copying.

In order to achieve the above object, the present invention provides an information processing system including a first information processing apparatus having a first communication port for transmitting and receiving data; a second information processing apparatus having a second communication port for transmitting and receiving data; and a communicating portion for executing bi-directional communication between the first communication port and the second communication port, 10 the information processing system comprising a utilizing portion for utilizing the communicating portion, for communication in a direction for which a first application program run on the first information processing apparatus sets the first communication port and the second communication port respectively as the sender and the destination of data; and a utilizing portion for utilizing the communicating portion, for communication in a direction for which a second application program run on the second information processing apparatus sets the 15 second communication port and the first communication port respectively as the sender and the destination of data.

The other features and objects than the above will become more apparent when reading the descriptions of 20 this specification with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

For the full understanding of the present invention and advantages thereof, the following description should 5 be referred to in conjunction with the accompanying drawings in which:

Fig. 1 is a block diagram showing the entire configuration of a storage device controller system according to an embodiment;

10 Fig. 2 is a diagram showing the layer configuration of a fiber channel protocol according to the embodiment;

Fig. 3 is a diagram showing the configuration of a data frame of the fiber channel protocol according to the embodiment;

15 Fig. 4 is a diagram showing the configuration of a header portion of the data frame of the fiber channel protocol according to the embodiment;

Fig. 5 is a block diagram illustrating data communication paths in the storage device controller 20 system according to the embodiment;

Fig. 6 is a diagram showing a path management table in the storage device controller system according to the embodiment;

25 Fig. 7 is a diagram showing an MCU unit path management table in the storage device controller system according to the embodiment;

Fig. 8 is a diagram showing an RCU unit path management

table in the storage device controller system according to the embodiment;

Fig. 9 is a diagram showing a path unit management table in the storage device controller system according to the embodiment;

Fig. 10 is a block diagram illustrating remote copying in the storage device controller system according to the embodiment;

Fig. 11 is a diagram showing a pair volume management table in the storage device controller system according to the embodiment; and

Fig. 12 is a diagram showing an M-VOL unit management table in the storage device controller system according to the embodiment.

15

DETAILED DESCRIPTION OF THE INVENTION

- Example of Overall Configuration -

An embodiment of the invention will now be described in greater detail with reference to the drawings.

20

Referring first to Fig. 1, a block diagram is shown which illustrates the entire configuration of a storage device controller system according to an embodiment.

25

An information processing apparatus A 100a is a computer comprising a CPU (Central Processing Unit), a memory etc. It receives an order from an operation terminal (not shown) connected with an LAN 300 and executes a predetermined program. This apparatus can be used as,

for example, a core computer in a large-scale computer system such as an ATM system of a bank and a seat reservation system for ticketing of airlines.

An information processing apparatus B 100b is also
5 a computer comprising a CPU, a memory etc. It is the computer that operates in place of the information processing apparatus A 100a when the information processing apparatus A 100a is disabled from operating by occurrence of a disaster etc. The information
10 processing apparatus B 100b is placed on a remote site far from the information processing apparatus A 100a and is connected with the information processing apparatus A 100a through the LAN 300.

The information processing apparatus A 100a is
15 connected with a storage apparatus A 200a through a data input/output link A 500a. The storage apparatus A 200a comprises an input/output control unit A 220a (a storage device controller) and a storage volume A (a storage device, M-VOL (Master Volume)) 210a and stores data interchanged
20 between the storage apparatus A 200a and the information processing apparatus A 100a.

The input/output control unit A 220a comprises a CPU, a memory etc. and controls input/output of data sent and received between the information processing apparatus
25 A 100a and the storage volume A 210a. Input/output of data between the information processing apparatus A 100a is carried out through a data input/output link A 500a

connected with a communication port A 221a.

The input/output control unit A 220a is also provided with another communication port A 221a which is connected with a remote copying link 400. The remote copying link 5 400 is a data transferring path connecting physically the storage apparatus A 200a and a storage apparatus B 200b.

In such a computer system, remote copying is carried out in the following procedure. That is, the ordinary 10 task is executed on the main site provided with the information processing apparatus A 100a and the storage apparatus A 200a. For writing data from the information processing apparatus A 100a to the storage apparatus A 200a, the input/output control unit A 220a transmits a 15 duplicate of the data through the remote copying link 400 to an input/output control unit B 220b provided to the storage apparatus B 200b. Then, the input/output control unit B 220b writes the data to a storage volume B (R-VOL (Remote Volume)) 210b.

20 In this way, remote copying causes both of the storage apparatus A 200a and the storage apparatus B 200b to have same data. It is possible to adapt such that the data is doubly written in the storage apparatus A 200a and the storage apparatus B 200b. Furthermore, it is also 25 possible to adapt such that a duplicate of data to be written to the storage apparatus A 200a is written to the storage apparatus B 200b. Yet furthermore, data

having been written in the storage apparatus A 200a is duplicated in the storage apparatus B 200b. By carrying out remote copying, it is possible to execute tasks on the remote site even when execution of the tasks is
5 impossible on the main site in case of a disaster.

Surely, it is also possible to remote-copy mutually on both of the main site and remote site. Adapting in this way, in case a disaster occurs to the system on one of the sites, tasks can be executed on the other site
10 while the main site and the remote site respectively execute their individual tasks in their normal operation,

It is possible to adapt such that the computer system on the remote site continues executing tasks as has been and the execution of tasks is returned to the system on
15 the main site when the computer system on the main site has been recovered.

In the former case, it is necessary to make settings for changing the direction in which duplicates of data is transmitted such that the site having been the remote
20 site becomes a main site and the site having been the main site functions as a remote site.

In the latter case, it is necessary to update the data on the main site by duplicating the data updated on the remote site after the occurrence of the failure
25 on the main site. Therefore, it is necessary to make settings such that the data is transmitted in a direction from the remote site to the main site.

In this way, it is necessary to make settings in advance as to where the data is transmitted from and to when carrying out remote copying. The contents of the settings are, for example, the settings of path of the 5 remote copying link 400. That is, they are settings defining the items such as which one of the storage apparatus A 200a and the storage apparatus B 200b is the sender of data and the destination of data, or which one of the communication ports 221a the data is transmitted 10 from and which one of communication ports 221b the data is transmitted to. Furthermore, it is necessary to make correspondence between specifying of a logic volume created logically on the storage area provided by the storage volume A 210a and data of which logic volume is 15 the subject of remote copying, and between specifying of a logic volume created logically on the storage area provided by the storage volume B 210b and which logic volume the data is to be written to. The contents of these settings will be described later.

20

- As to Communication Protocol -

Communication through the remote copying link 400 is carried out according to a predetermined communication protocol. Among the many fiber channel protocols defined 25 by ANSI (American National Standard Institute), a protocol called FC-SB-2 can be used for remote copying. The details of it is described in "FIBER CHANNEL PHYSICAL AND SIGNALING

INTERFACE (FC-PH) Rev. 4.3" of the ANSI standards but FC-SB-2 is formed by mapping a protocol used for main frames on FC-PH by which the physical layer of a fiber channel protocol is defined. Fig. 2 shows the layer 5 configuration of this fiber channel protocol.

Among these, FC-SB-2 is a communication program characterized by the following items.

First, bi-directional communication can be available with it. Thereby, it is possible for 10 communication ports on both ends on a physical link can simultaneously transmit data. In contrast, in FC-AL being a communication protocol for open systems, it is necessary for any one of communication ports to obtain at any time the right to transmit in order to be the initiator, 15 and to carry out transferring of data to a target when the data is transferred. Therefore, during this operation, other communication ports can not be the initiator and can not carry out data transfer. Another characteristic is that multiple I/O (Input/Output) 20 operations that are impossible in the conventional main frame protocols can be available conforming to most of the conventional main frame protocols. Thereby, multiple I/O process to a plurality of devices through one same physical link that is conventionally impossible can be 25 available.

Fig. 3 shows the configuration of a data frame 700 of the fiber channel protocol according to the embodiment.

"SOF (Start Of Frame) delimiter" is a character indicating the start of a data frame. "EOF (End Of Frame) delimiter" is a character indicating the end of a data frame. "Data Field" is an area for storing data to be transmitted. "FC-PH-CRC" is where a check code of a frame data is stored in. Information indicating the type of ULPS (Upper Level Protocols) among the protocol layers shown in Fig. 2 and information indicating the initiator having issued the data frame etc. are stored in "Frame Header 710".

Fig. 4 shows the configuration of a frame header 710. "RCTL" indicates the protocol of the data frame and the type of the data frame. The address of a destination port is stored in "D-ID (Destination ID)". The address of a sender port is stored in "S-ID (Source ID)". A more specific code indicating the type of the frame defined by RCTL is stored in "TYPE". When the ULPS (Upper Level Protocols) are FC-SB-2, information indicating the initiator having issued the data frame is stored. Bits for managing and controlling I/O as a logical lump is stored in "F-CTL". "SEQ-ID" indicates an ID given to a frame sequence. "SEQ-ONT" indicates an ID of a frame constituting one frame sequence. Information indicating whether the data frame is optical or not is stored in "DF-CTL". "OX-ID" and "RX-ID" indicate respectively the reference numbers given to the issuer and the receiver of I/O. Various pieces of information are stored in

"Parameter" in response to the contents of RCTL.

By using such communication protocol, it is possible to carry out data transfers for remote copying.

5 - As to Logic Paths -

Next, Fig. 5 shows a block diagram for illustrating a data communication in a storage device controller system according to the embodiment.

The storage apparatus A 200a and the storage apparatus 10 B 200b are connected mutually through the physical path (a remote copying link) 400. The storage apparatus A 200a consists of n+1 CU (Control Unit) images 230a. A CU image 230a is a storage apparatus provided virtually on the storage apparatus A 200a. Thereby, it appears as if each 15 of n+1 storage apparatuses is connected respectively to each of CU images 230a from the information processing apparatus A 100a shown in Fig. 1. The storage apparatus B 200b is also consists of n+1 CU images 230b.

The connection between the storage apparatus A 200a 20 and the storage apparatus B 200b through the remote copying link 400 is carried out by setting up a logical path (logic path) 410 for each of the CU images 230. The following logic paths are shown being set up in Fig. 5, That is, from the CU image #0 of the storage apparatus A 200a, 25 logic paths toward the CU image #0 and CU image #1 of the storage apparatus B 200b are set up. From the CU image #1 of the storage apparatus A 200a, a logic path toward

the CU image #1 of the storage apparatus B 200b is set up. From the CU image #0 of the storage apparatus B 200b, a logic path toward the CU image #0 of the storage apparatus A 200a is set up. From the CU image #1 of the storage apparatus B 200b, a logic path toward the CU image #1 of the storage apparatus A 200a is set up.

The data transfer between the storage apparatus A 200a and the storage apparatus B 200b is carried out by setting up logic paths in advance as described above.

10

- Setting up of Logic Paths -

Next, Figs. 6-9 show tables for carrying out the setting up of the above logic paths (communication direction defining information). Creation of such tables is carried out by a program (application program) run on the storage apparatus A 200a. First, Fig. 6 shows a path management table 800 in the storage device controller system according to the embodiment.

The path management table 800 is an aggregation of MCUs (Master CU) unit path management table 810 provided to each of the CU images 230. MCU herein means a CU of the storage apparatus on the main site.

Fig. 7 shows a MCU unit path management table 810. The MCU unit path management table 810 is an aggregation of RCU valid flag information 811 and n+1 RCU (Remote CU) unit path management tables 820. RCU herein means a CU of the storage apparatus on the remote site.

The RCU valid flag information 811 is a bitmap for each RCU and, when it is "1", it indicates that the information is paired with the RCU and, when it is "0", it indicates that the information is not paired with the
5 RCU.

As shown in Fig. 8, the RCU unit path management table 820 is an aggregation of a path number valid table 821 and $m+1$ path unit management tables 830. The path number valid table 821 is a bitmap indicating whether each of
10 $m+1$ logic paths capable of being formed between the RCU is respectively valid or invalid. When it is "1", it indicates that the logic path is valid and, when it is "0", it indicates that the logic path is invalid.

As shown in Fig. 9, the path unit management table 830 is an aggregation of the path number 831, the sender port number 832, the destination port number 833, path direction information 834, path status information 835 and a path formation time 836.

The path direction information 834 is in a bitmap form. For example, in the case of information of four (4) bytes, when bit zero (0) is "1", it indicates that a path is formed with the MCU as the initiator and when bit zero (0) is "0", it indicates that the MCU is not the initiator. When bit one (1) is "1", it indicates that
20 a path is formed with RCU as the initiator and the bit one (1) is "0", it indicates the RCU is not the initiator.
25 When paths are formed in both of the directions, these

bits are summed.

The path status information 835 indicates the status of the path. When the information is "0x01", it indicates the path is not formed, when it is "0x02", it indicates 5 the path is being formed, when it is "0x03", it indicates the path has been formed, when it is "0x04", it indicates the path is being suspended and, when it is "0x05", it indicates the path is being re-formed. "0x" herein indicates that the digit coming after this is a hexadecimal 10 digit. For example, if a failure occurs on the main site when the path has been formed and a status where the pair is formed can not be maintained temporarily, the status of the path is shifted to the status where the path is suspended. Then, if the path is formed from the main site 15 after the main site has been recovered, the status of the path is shifted to the status where the path has been formed through the status where the path is being re-formed.

The path formation time 836 stores the system time 20 when the path is formed.

In the storage device controller system according to the embodiment, using the above tables, the logic path 410 is formed on the remote copying link 400 being the data transfer path between the storage apparatus A 200a 25 and the storage apparatus B 200b and the sender and the destination of data for carrying out remote copying are managed. Thereby, it is possible to carry out easily the

setting up of a path in the opposite direction set up when the main site has been recovered.

That is, on the remote site, by re-constituting tables such that the RCU numbers and MCU numbers in the above 5 tables are exchanged, it is possible to form a logic path in the opposite direction. Such re-configuration of the tables is carried out with a program (application program) run on the storage apparatus B 200b. Surely, when the re-configuration of the tables is carried out, the sender 10 port number 832 and the destination port number 833 in the path unit management table are exchanged with each other.

In this way, in the storage device controller system according to the embodiment, it is possible to carry out 15 easily the setting up of a path on the remote site by using the above tables.

Furthermore, in the storage device controller system according to the embodiment, the information indicating which storage apparatus is the initiator for forming a 20 path in terms of the logic path 410 formed on the physically same remote copying link 400 can be managed. Thereby, it is possible to carry out bi-directional communication on the physically same remote copying link 400 from two (2) storage apparatuses. Therefore, it is possible to 25 effectively utilize the physical path resource between the two (2) storage apparatuses.

Yet furthermore, the management of settings of the

data transfer paths according to the embodiment can be applied not only to the settings for data transfer paths for remote copying but also to the settings of the data transfer paths in common information processing systems.

- 5 carrying out data transfer using communication protocols in which bi-directional communication is possible. Similarly, it can be applied to the settings of the data transfer path between the information processing apparatus A 100a and the storage apparatus A 200a.

10

- As to Remote Copying -

Next, a block diagram for illustrating remote copying carried out in the storage device controller system according to the embodiment is shown in Fig. 10.

- 15 The storage apparatus A 200a and the storage apparatus B 200b are connected mutually by the physical path (remote copying link) 400. The storage apparatus A 200a consists of n+1 CU images 230a. Similarly, the storage apparatus B 200b consists of n+1 CU images 230b.

- 20 The connection between the storage apparatus A 200a and the storage apparatus B 200b through the remote copying link 400 is carried out by setting up the logical path (logic path) 410 for each of the CU images 230. The settings for the logic path 410 is same as in Fig. 5,

- 25 In Fig. 10, the CU image #0 of the storage apparatus A 200a manages M-VOL 210a. The CU image #0 of the storage apparatus B 200b manages R-VOL 210b. The settings for

remote copying have been carried out such that, when new data is written to the M-VOL 210a, the duplicate of the data is written to the R-VOL 210b.

5 - Settings for Remote Copying-

Tables for making settings for such remote copying are shown in Fig. 11 and Fig. 12.

A pair volume management table 900 shown in Fig. 11 is an aggregation of k+1 M-VOL unit management tables 910 shown in Fig. 12. The M-VOL unit management table 910 is created for each of M-VOL 210a created in the storage apparatus A 200a and stores information indicating whether pair forming is present or not 911, and showing the R-VOL number 912, the sender port number 913, the destination port number 914 and a path number 915. The creation of the pair volume management table 900 is carried out with a program (application program) run on the storage apparatus A 200a.

The information indicating whether the pair forming is present or not 911 indicates whether the M-VOL 210a is targeted for remote copying or not. When it is "1", the M-VOL 210a is targeted for remote copying and, when it is "0", the M-VOL 210a is not targeted for remote copying. When the M-VOL 210a is targeted for remote copying, a duplicate of the data written in the M-VOL is transmitted to the remote site. R-VOL number 912 indicates the number of a storage volume for storing data for remote copying

transmitted to the remote site. The sender port number 913 indicates the communication port 221a number on the main site. The destination port number 914 indicates the communication port 221b number on the remote site. The 5 path number 915 indicates the logic path number for transmitting the data for remote copying.

In the storage device controller system according to the embodiment, using the above tables, the sender storage volume (M-VOL) and destination storage volume 10 (R-VOL) can be managed when remote copying is carried out. Thereby, it is possible to easily make settings for remote copying in the opposite direction set up when the main site has been recovered from a disaster.

That is, it is possible to make settings for remote 15 copying in the opposite direction by re-constituting the tables such that the R-VOL number and M-VOL number are exchanged with each other in the above tables on the remote site. Such re-configuration of the tables can be carried out with a program (application program) run on the storage 20 apparatus B 200b. Surely, when the re-configuration of the tables is carried out, the sender port number 913 and the destination port number 914 in the M-VOL unit management table 910 are exchanged with each other.

In this way, in the storage device controller system 25 according to the embodiment, it is possible to make settings easily for the path on the remote site.

The embodiment has been described, however, the above

embodiment is intended only to facilitate understanding
the invention and not intended to limit the interpretation
of the invention. Various changes and modifications of
the invention can be made without departing from the scope
5 and spirit of the invention and the invention includes
its equivalents.

The invention can thus provide an information system,
a storage system, a storage device controller and a
program.

10 While illustrative and presently preferred
embodiments of the present invention have been described
in detail herein, it is to be understood that these
embodiments could variously be modified, permuted or
altered without departing from the spirit and scope of
15 the invention defined by the appended claims.